

Recursive Least Squares and Multi-innovation Stochastic Gradient Parameter Estimation Methods for Signal Modeling

Ling Xu^{1,2} · Feng Ding¹

Received: 25 February 2016 / Revised: 23 July 2016 / Accepted: 25 July 2016

Published online: 11 August 2016

© Springer Science+Business Media New York 2016

Abstract The sine signals are widely used in signal processing, communication technology, system performance analysis and system identification. Many periodic signals can be transformed into the sum of different harmonic sine signals by using the Fourier expansion. This paper studies the parameter estimation problem for the sine combination signals and periodic signals. In order to perform the online parameter estimation, the stochastic gradient algorithm is derived according to the gradient optimization principle. On this basis, the multi-innovation stochastic gradient parameter estimation method is presented by expanding the scalar innovation into the innovation vector for the aim of improving the estimation accuracy. Moreover, in order to enhance the stabilization of the parameter estimation method, the recursive least squares algorithm is derived by means of the trigonometric function expansion. Finally, some simulation examples are provided to show and compare the performance of the proposed approaches.

Keywords Signal modeling · Recursive algorithm · Fourier series · Parameter estimation · Least squares · Sine signal

✉ Ling Xu
lingxu0848@163.com

Feng Ding
fding@jiangnan.edu.cn

¹ School of Internet of Things Engineering, Jiangnan University, Wuxi 214122, People's Republic of China

² School of Internet of Things Technology, Wuxi Vocational Institute of Commerce, Wuxi 214153, People's Republic of China

1 Introduction

The sine signal is a single frequency wave and is used widely in the communication technology, signal processing and filtering [25] and system identification [2,37]. Many period signals under certain conditions can be decomposed into the sine combinations with different frequencies, amplitudes and phases. The signal modeling is estimating the characteristic parameters of the signals from the measured data. The sine-wave parameter estimation problems have been received much attentions. For example, Belega et al. [3] studied the accuracy of the sine-wave parameter estimation by means of the windowed three-parameter sine-fitting algorithm; Chen et al. [5] studied the multi-harmonic fitting algorithm based on four parameters sine fitting to improve the global convergence; Li et al. [19] derived a gradient-based iterative identification algorithm for estimating parameters of the signal model with known and unknown frequencies. Some of these methods are based on the statistic analysis. This paper considers the optimization algorithm for estimating the signal parameters.

The mathematical model is the basics of controller design [36,39]. Many identification algorithms can estimate system parameters and system models [15,26]. In general, the identification algorithm is derived by defining and minimizing a cost function [27–32,35]. The parameters to be estimated are defined as the parameter vector. Then, the parameter estimates can be obtained by means of optimization algorithms. The identification algorithms have been used widely in industrial robot, signal processing and network communication [8,41]; Guo et al. [14] investigated the recursive identification method for finite impulse response systems with binary-valued outputs and communication channels; Janot et al. [16] addressed a revised Durbin–Wu–Hausman test for the industrial robot identification; Zhao et al. [43] studied the multi-frequency identification algorithm to identify the amplitude and phase of a multi-frequency signal. This paper studies the application of the identification algorithm to the signal modeling.

In system identification, some identification algorithms focus on reducing the computation load and enhancing the accuracy [40]. Since the gradient optimization method only needs computing the first-order derivation, the gradient identification algorithm has low computation load [10]. However, the gradient algorithm has low computation accuracy. Many improved gradient algorithms have been proposed for enhancing the computation accuracy. Andrei [1] illustrated an adaptive conjugate gradient algorithm for large-scale unconstrained optimization by minimizing the quadratic approximation of the objective function at the current point; Deng et al. [7] developed a three-term conjugate gradient algorithm for solving large-scale unconstrained optimization problems by rectifying the steepest descent direction with the difference between the current iterative points and the gradients; Necoara et al. [24] devised a fully distributed dual gradient method based on a weighted step size and analyzed the convergence rate. Although these improved algorithms can polish up the convergence rate and the estimation accuracy, the computation load is heavy.

The innovation is the useful information that can improve the parameter estimation accuracy. It can promote the convergence of the algorithms during the recursive

process. In order to enhance the estimation accuracy by using more innovation, the multi-innovation theory is used widely in the system identification. Mao et al. [23] studied a data filtering-based multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems; Zhang et al. [42] considered a multi-innovation auto-constructed least squares identification method for 4 freedom ship maneuvering identification modeling; in this paper, the multi-innovation method is expanded into the signal modeling for the sine-wave or periodic signals.

In general, the identification methods are divided into the online identification and the off-line identification. The iterative identification methods are used to the off-line identification [12,38]. In view of the online identification, Li et al. [22] studied a parallel adaptive self-tuning recursive least squares algorithm for the time-varying system; Ding et al. [11] developed a recursive least squares parameter identification algorithms for nonlinear systems; Ding et al. [13] studied a recursive least squares parameter estimation method for a class of output nonlinear systems based on the model decomposition; Ding et al. [9] presented a least squares algorithm for a dual-rate state space system with time delay. Considering the advantages of the least squares algorithm, a recursive least squares parameter estimation algorithm is derived to estimate the parameters of a periodic signal.

The major contributions of the work in this paper are listed in the following.

- This paper studies the problem of the parameter estimation. The proposed methods can be used not only for the sine combination signals but also for other periodic signals. These parameter estimation methods can be used in signal processing and signal modeling.
- On the basis of the gradient searching, a stochastic gradient (SG) parameter estimation algorithm is presented. In order to improve the estimation accuracy and the convergence rate, a multi-innovation stochastic gradient (MISG) parameter estimation algorithm is presented by means of expanding the scalar innovation into the innovation vector.
- For the purpose of enhancing the algorithm stabilization and the parameter estimation accuracy, a recursive least squares (RLS) algorithm is derived using the trigonometric function expansion. This technique transforms the nonlinear optimization into the linear optimization. Therefore, the algorithm stabilization is improved significantly.

The rest of this paper is organized in the following. Section 2 derives the SG method. Section 3 deduces the MISG algorithm. Section 4 gives the RLS parameter estimation algorithm. Section 5 provides some examples to illustrate and compare the effectiveness of the proposed parameter estimation methods. Section 6 draws some concluding remarks.

2 Stochastic Gradient Method

Let us introduce some notation.

Symbol	Meaning
X^T	The transpose of the vector or matrix X
$\hat{\theta}(k)$	The estimate of θ at recursion k
$A := X$	I.e., $X := A$, X is defined by A
$\text{tr}[X]$	The trace of the square matrix X
$\ X\ $	$\ X\ ^2 := \text{tr}[XX^T]$
$e(k)$	The innovation scalar at recursion k
$E(p, k)$	The innovation vector, i.e., the multi-innovation at recursion k
$\varphi(k)$	The information vector at recursion k
$\Phi(p, k)$	The information matrix at recursion k

Consider the combination sine signals with different frequencies, phases and amplitudes:

$$y(t) = a_1 \sin(\omega_1 t + \phi_1) + a_2 \sin(\omega_2 t + \phi_2) + \dots + a_n \sin(\omega_n t + \phi_n), \quad (1)$$

where a_1, a_2, \dots, a_n are the amplitudes, $\omega_1, \omega_2, \dots, \omega_n$ are the frequencies and $\phi_1, \phi_2, \dots, \phi_n$ are the phases. These parameters are the characteristic parameters of the combination sine signals. The goal is to estimate these parameters by means of presenting new identification methods.

Suppose that the frequencies of the combination sine signals are known, then the phases and amplitudes are to be identified. Define the parameter vector

$$\theta := [a_1, a_2, \dots, a_n, \phi_1, \phi_2, \dots, \phi_n]^T \in \mathbb{R}^{2n}.$$

In the identification test, assume that the sampling period is h and the sampling time is $t_k := kh$. The measured data are represented as $\{t_k, y(t_k)\}$. Let $y(k) := y(t_k)$ for simplification. Define the difference between the observation output and the model output:

$$\varepsilon(k) := y(k) - \sum_{i=1}^n a_i \sin(\omega_i kh + \phi_i).$$

Then, define the cost function

$$J(\theta) := \frac{1}{2} \varepsilon^2(k).$$

Taking the first-order derivative of $J(\theta)$ with respect to θ gives

$$\begin{aligned} \text{grad}[J(\theta)] &:= \frac{\partial J(\theta)}{\partial \theta} \\ &= \left[\frac{\partial J(\theta)}{\partial a_1}, \frac{\partial J(\theta)}{\partial a_2}, \dots, \frac{\partial J(\theta)}{\partial a_n}, \frac{\partial J(\theta)}{\partial \phi_1}, \frac{\partial J(\theta)}{\partial \phi_2}, \dots, \frac{\partial J(\theta)}{\partial \phi_n} \right]^T \in \mathbb{R}^{2n}, \end{aligned}$$

$$\begin{aligned}\frac{\partial J(\boldsymbol{\theta})}{\partial a_i} &= -\sin(\omega_i kh + \phi_i)\varepsilon(k), \\ \frac{\partial J(\boldsymbol{\theta})}{\partial \phi_i} &= -a_i \cos(\omega_i kh + \phi_i)\varepsilon(k), \quad i = 1, 2, \dots, n.\end{aligned}$$

Define the information vector

$$\boldsymbol{\varphi}(k) := [\sin(\omega_1 kh + \phi_1), \dots, \sin(\omega_n kh + \phi_n), a_1 \cos(\omega_1 kh + \phi_1), \dots, a_n \cos(\omega_n kh + \phi_n)]^\top \in \mathbb{R}^{2n}.$$

Let $k = 1, 2, \dots$ be a recursive variable and let $\hat{\boldsymbol{\theta}}(k)$ be the estimate of $\boldsymbol{\theta}$ at recursion k . Utilizing the gradient search and minimizing the cost function $J(\boldsymbol{\theta})$, we have the SG parameter estimation algorithm:

$$\hat{\boldsymbol{\theta}}(k+1) = \hat{\boldsymbol{\theta}}(k) + \frac{\hat{\boldsymbol{\varphi}}(k)}{r(k+1)} e(k), \quad (2)$$

$$e(k) = y(k) - \sum_{i=1}^n \hat{a}_i(k) \sin(\omega_i kh + \hat{\phi}_i(k)), \quad (3)$$

$$\hat{\boldsymbol{\varphi}}(k) = [\sin(kh\omega_1 + \hat{\phi}_1(k)), \dots, \sin(kh\omega_n + \hat{\phi}_n(k)), \hat{a}_1(k) \cos(\omega_1 kh + \hat{\phi}_1(k)), \dots, \hat{a}_n(k) \cos(\omega_n kh + \hat{\phi}_n(k))]^\top, \quad (4)$$

$$r(k+1) = r(k) + \|\hat{\boldsymbol{\varphi}}(k)\|^2, \quad r(0) = 1. \quad (5)$$

The steps of computing the parameter estimate $\hat{\boldsymbol{\theta}}(k+1)$ using the SG method are as follows.

1. To initiate: let $k = 0$, preset the recursive length L and let $\hat{\boldsymbol{\theta}}(0)$ be an arbitrary small real vector.
2. Collect the measured data $y(k)$.
3. Compute $\hat{\boldsymbol{\varphi}}(k)$ using (4) and compute $r(k+1)$ using (5).
4. Compute $e(k)$ using (3).
5. Update $\hat{\boldsymbol{\theta}}(k+1)$ using (2), if $k = L$, then terminate the recursive procedure; otherwise, $k := k + 1$, go to Step 2.

3 The Multi-innovation Stochastic Gradient Algorithm

For the SG algorithm, the SG algorithm has low computation accuracy. In order to enhance the computation precision, more measurement data are utilized to the algorithm at each recursion. The dynamical window data scheme is adopted to derive the MISG algorithm. The dynamical window data are a batch data, and the data length is p . In the SG algorithm, $e(k)$ is called the scalar innovation. This innovation can promote the algorithm estimation accuracy.

Consider that the dynamical window data with length p are $y(k), y(k - 1), \dots, y(k - p + 1)$. Expand the scalar innovation $e(k)$ into the innovation vector

$$\mathbf{E}(p, k) := \begin{bmatrix} y(k) - \sum_{i=1}^n \hat{a}_i(k) \sin(\omega_i kh + \hat{\phi}_i(k)) \\ y(k - 1) - \sum_{i=1}^n \hat{a}_i(k) \sin((kh - h)\omega_i + \hat{\phi}_i(k)) \\ \vdots \\ y(k - p + 1) - \sum_{i=1}^n \hat{a}_i(k) \sin((kh - ph + h)\omega_i + \hat{\phi}_i(k)) \end{bmatrix} \in \mathbb{R}^p,$$

where $\hat{a}_i(k)$ denotes the estimate of a_i and $\hat{\phi}_i(k)$ denotes the estimate of ϕ_i at time $t = kh$. Expand the information vector into the information matrix

$$\Phi(p, k) := [\boldsymbol{\varphi}(k), \boldsymbol{\varphi}(k - 1), \dots, \boldsymbol{\varphi}(k - p + 1)] \in \mathbb{R}^{2n \times p}.$$

According to the gradient searching, the MISG parameter estimation algorithm is listed in the following:

$$\hat{\boldsymbol{\theta}}(k + 1) = \hat{\boldsymbol{\theta}}(k) - \frac{\hat{\Phi}(p, k)}{r(k + 1)} \mathbf{E}(p, k), \quad (6)$$

$$\mathbf{E}(p, k) = [e(k), e(k - 1), \dots, e(k - p + 1)]^T, \quad (7)$$

$$e(k - i) = y(k - i) - \sum_{j=1}^n \hat{a}_j(k) \sin((k - i)h\omega_j + \hat{\phi}_j(k)), \quad i = 0, 1, \dots, p - 1, \quad (8)$$

$$\Phi(p, k) = [\hat{\boldsymbol{\varphi}}(k), \hat{\boldsymbol{\varphi}}(k - 1), \dots, \hat{\boldsymbol{\varphi}}(k - p + 1)], \quad (9)$$

$$\hat{\boldsymbol{\varphi}}(k - i) = [-\sin((k - i)h\omega_1 + \hat{\phi}_1(k)), \dots, -\sin((k - i)h\omega_n + \hat{\phi}_n(k)), \\ -\hat{a}_1(k) \cos((k - i)h\omega_1 + \hat{\phi}_1(k)), \dots, -\hat{a}_n(k) \cos((k - i)h\omega_n \\ + \hat{\phi}_n(k))]^T, \quad i = 0, 1, \dots, p - 1, \quad (10)$$

$$r(k + 1) = r(k) + \|\Phi(p, k)\|^2, \quad r(0) = 1. \quad (11)$$

The steps of computing the parameter estimate $\hat{\boldsymbol{\theta}}(k + 1)$ using the MISG method are as follows.

1. To initiate: preset the recursive length L and the innovation length p ; let $\hat{\boldsymbol{\theta}}(0)$ be an arbitrary small real vector.
2. Collect measured data $y(k)$; compute $\hat{\boldsymbol{\varphi}}(k - i)$ using (10); form $\Phi(p, k)$ using (9).
3. Compute $e(k - i)$ using (8) and form $\mathbf{E}(p, k)$ using (7).
4. Compute $r(k + 1)$ using (11).
5. Update the parameter estimate $\hat{\boldsymbol{\theta}}(k + 1)$ using (6), if $k = L$, then terminate the recursive procedure; otherwise $k := k + 1$, go to Step 2.

4 The Recursive Least Squares Algorithm

The least squares optimization method is widely used in system identification. This paper expands this method to the signal modeling. It is obvious that the sine combi-

nation signal is a nonlinear function with respect to the parameters to be estimated. In order to derive the recursive least squares algorithm to estimate the parameters of the combination sine signal, rewriting the sine combination signal in (1) gives

$$y(t) = a_1 \cos \phi_1 \sin \omega_1 t + a_1 \sin \phi_1 \cos \omega_1 t + a_2 \cos \phi_2 \sin \omega_2 t + a_2 \sin \phi_2 \cos \omega_2 t + \cdots + a_n \cos \phi_n \sin \omega_n t + a_n \sin \phi_n \cos \omega_n t. \quad (12)$$

Let $c_i := a_i \cos \phi_i$, $d_i := a_i \sin \phi_i$, thus $y(t)$ can be expressed as

$$y(t) = \sum_{i=1}^n c_i \sin \omega_i t + d_i \cos \omega_i t.$$

Define the parameter vector

$$\boldsymbol{\theta} := [c_1, d_1, c_2, d_2, \dots, c_n, d_n]^T \in \mathbb{R}^{2n}.$$

In the identification test, the sampling time is $t_k := kh$ and the measured data length is L . The observation output data are $y(k) := y(kh)$, $k = 1, 2, \dots$

Define the information vector

$$\boldsymbol{\varphi}(k) := [\sin \omega_1 kh, \cos \omega_1 kh, \sin \omega_2 kh, \cos \omega_2 kh, \dots, \sin \omega_n kh, \cos \omega_n kh]^T \in \mathbb{R}^{2n}.$$

Define the criterion function

$$J(\boldsymbol{\theta}) := \frac{1}{2} \sum_{j=1}^k e^2(j),$$

where $e(j) := y(j) - \boldsymbol{\varphi}^T(j)\boldsymbol{\theta}$.

Define the stack output vector \mathbf{Y}_k and the stack information vector $\boldsymbol{\Phi}_k$, respectively,

$$\mathbf{Y}_k := \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(k) \end{bmatrix} \in \mathbb{R}^k, \quad \boldsymbol{\Phi}_k := \begin{bmatrix} \boldsymbol{\varphi}^T(1) \\ \boldsymbol{\varphi}^T(2) \\ \vdots \\ \boldsymbol{\varphi}^T(k) \end{bmatrix} \in \mathbb{R}^{k \times 2n}.$$

Then, the criterion function can be rewritten as

$$J(\boldsymbol{\theta}) = \frac{1}{2} \|\mathbf{Y}_k - \boldsymbol{\Phi}_k \boldsymbol{\theta}\|^2.$$

Minimizing the criterion function $J(\boldsymbol{\theta})$ gives the least squares parameter estimate

$$\hat{\boldsymbol{\theta}}(k) = (\boldsymbol{\Phi}_k^T \boldsymbol{\Phi}_k)^{-1} \boldsymbol{\Phi}_k^T \mathbf{Y}_k. \quad (13)$$

Obviously, the least squares estimate $\hat{\boldsymbol{\theta}}(k)$ involves computing the inverse matrix $(\boldsymbol{\Phi}_k^T \boldsymbol{\Phi}_k)^{-1}$. In order to avoid computing the inverse matrix, this paper develops a RLS algorithm to estimate the parameters of the combination sine signal for the online estimation. Define a matrix

$$\mathbf{P}^{-1}(k) := \boldsymbol{\Phi}_k^T \boldsymbol{\Phi}_k = \sum_{j=1}^k \boldsymbol{\varphi}(j) \boldsymbol{\varphi}^T(j). \quad (14)$$

Rewriting (14), we have

$$\mathbf{P}^{-1}(k) = \sum_{j=1}^{k-1} \boldsymbol{\varphi}(j) \boldsymbol{\varphi}^T(j) + \boldsymbol{\varphi}(k) \boldsymbol{\varphi}^T(k). \quad (15)$$

Uteriorly, Eq. (15) can be represented as

$$\mathbf{P}^{-1}(k) = \mathbf{P}^{-1}(k-1) + \boldsymbol{\varphi}(k) \boldsymbol{\varphi}^T(k) = \mathbf{P}^{-1}(0) + \sum_{j=1}^k \boldsymbol{\varphi}(j) \boldsymbol{\varphi}^T(j),$$

where $\mathbf{P}^{-1}(0) = p_0 \mathbf{I} > 0$, $p_0 = 10^6$.

Because $\mathbf{Y}_k := \begin{bmatrix} \mathbf{Y}_{k-1} \\ y(k) \end{bmatrix} \in \mathbb{R}^k$, $\boldsymbol{\Phi}_k := \begin{bmatrix} \boldsymbol{\Phi}_{k-1} \\ \boldsymbol{\varphi}^T(k) \end{bmatrix} \in \mathbb{R}^{k \times 2n}$, Eq. (13) can be represented as

$$\begin{aligned} \hat{\boldsymbol{\theta}}(k) &= \mathbf{P}(k) \boldsymbol{\Phi}_k^T \mathbf{Y}_k \\ &= \mathbf{P}(k) \begin{bmatrix} \boldsymbol{\Phi}_{k-1} \\ \boldsymbol{\varphi}^T(k) \end{bmatrix}^T \begin{bmatrix} \mathbf{Y}_{k-1} \\ y(k) \end{bmatrix} \\ &= \mathbf{P}(k) [\boldsymbol{\Phi}_{k-1} \mathbf{Y}_{k-1} + \boldsymbol{\varphi}(k) y(k)] \\ &= \mathbf{P}(k) [\mathbf{P}^{-1}(k-1) \mathbf{P}(k-1) \boldsymbol{\Phi}_{k-1} \mathbf{Y}_{k-1} + \boldsymbol{\varphi}(k) y(k)] \\ &= \mathbf{P}(k) [\mathbf{P}^{-1}(k-1) \hat{\boldsymbol{\theta}}(k-1) + \boldsymbol{\varphi}(k) y(k)] \\ &= \mathbf{P}(k) [\mathbf{P}^{-1}(k) - \boldsymbol{\varphi}(k) \boldsymbol{\varphi}^T(k)] \hat{\boldsymbol{\theta}}(k-1) + \mathbf{P}(k) \boldsymbol{\varphi}(k) y(k) \\ &= \hat{\boldsymbol{\theta}}(k-1) + \mathbf{P}(k) \boldsymbol{\varphi}(k) [y(k) - \boldsymbol{\varphi}^T(k) \hat{\boldsymbol{\theta}}(k-1)]. \end{aligned}$$

In order to avoid computing the inverse matrix $\mathbf{P}^{-1}(k)$, we use the matrix inverse forum $(\mathbf{A} + \mathbf{BC})^{-1} = \mathbf{A}^{-1} \mathbf{B} (\mathbf{I} + \mathbf{CA}^{-1} \mathbf{B})^{-1} \mathbf{CA}^{-1}$ to $\mathbf{P}^{-1}(k) = \mathbf{P}^{-1}(k-1) + \boldsymbol{\varphi}(k) \boldsymbol{\varphi}^T(k)$, and have

$$\mathbf{P}(k) = \mathbf{P}(k-1) - \frac{\mathbf{P}(k-1) \boldsymbol{\varphi}(k) \boldsymbol{\varphi}^{-1}(k) \mathbf{P}(k-1)}{1 + \boldsymbol{\varphi}(k) \mathbf{P}(k-1) \boldsymbol{\varphi}(k)}. \quad (16)$$

Introducing the gain vector $\mathbf{L}(k) := \mathbf{P}(k) \boldsymbol{\varphi}(k)$ and multiplying by $\boldsymbol{\varphi}(k)$ both side of (16), we have

$$\begin{aligned} \mathbf{P}(k)\boldsymbol{\varphi}(k) &= \left[\mathbf{P}(k-1) - \frac{\mathbf{P}(k-1)\boldsymbol{\varphi}(k)\boldsymbol{\varphi}^{-1}(k)\mathbf{P}(k-1)}{1 + \boldsymbol{\varphi}(k)\mathbf{P}(k-1)\boldsymbol{\varphi}(k)} \right] \boldsymbol{\varphi}(k) \\ &= \mathbf{P}(k-1)\boldsymbol{\varphi}(k) \left[1 - \frac{\boldsymbol{\varphi}(k)\mathbf{P}(k-1)\boldsymbol{\varphi}(k)}{1 + \boldsymbol{\varphi}(k)\mathbf{P}(k-1)\boldsymbol{\varphi}(k)} \right]. \end{aligned}$$

From the above analysis, we obtain the following equation

$$\mathbf{L}(k) = \mathbf{P}(k-1)\boldsymbol{\varphi}(k) \left[1 - \frac{\boldsymbol{\varphi}(k)\mathbf{P}(k-1)\boldsymbol{\varphi}(k)}{1 + \boldsymbol{\varphi}(k)\mathbf{P}(k-1)\boldsymbol{\varphi}(k)} \right].$$

However, the parameters in the parameter vector $\boldsymbol{\theta}$ are not the characteristic parameters of the combination sine signal. According to the previous definition $c_i := a_i \cos \phi_i$, $d_i := a_i \sin \phi_i$, the estimates of the characteristic parameters a_i and ϕ_i are computed by

$$\hat{a}_i(k) = \sqrt{\hat{c}_i^2(k) + \hat{d}_i^2(k)}, \quad \hat{\phi}_i(k) = \arctan \frac{\hat{d}_i(k)}{\hat{c}_i(k)}.$$

Let $\hat{\boldsymbol{\theta}}_r(k) := [\hat{a}_1(k), \dots, \hat{a}_n(k), \hat{\phi}_1(k), \dots, \hat{\phi}_n(k)]^T \in \mathbb{R}^{2n}$ denote the estimates of the characteristic parameters of the combination sine signal.

Finally, we obtain the RLS algorithm:

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + \mathbf{L}(k)[y(k) - \boldsymbol{\varphi}(k)\hat{\boldsymbol{\theta}}(k-1)], \quad (17)$$

$$\mathbf{L}(k) = \frac{\mathbf{P}(k-1)\boldsymbol{\varphi}(k)}{1 + \boldsymbol{\varphi}^T(k)[\mathbf{P}(k-1)\boldsymbol{\varphi}(k)]}, \quad (18)$$

$$\mathbf{P}(k) = \mathbf{P}(k-1) - \mathbf{L}(k)[\mathbf{P}(k-1)\boldsymbol{\varphi}(k)]^T, \quad \mathbf{P}(0) = p_0\mathbf{I}, \quad (19)$$

$$\boldsymbol{\varphi}(k) = [\sin \omega_1 kh, \cos \omega_1 kh, \sin \omega_2 kh, \cos \omega_2 kh, \dots, \sin \omega_n kh, \cos \omega_n kh]^T, \quad (20)$$

$$\hat{a}_i(k) = \sqrt{\hat{c}_i^2(k) + \hat{d}_i^2(k)}, \quad \hat{\phi}_i(k) = \arctan \frac{\hat{d}_i(k)}{\hat{c}_i(k)}, \quad i = 1, 2, \dots, n, \quad (21)$$

$$\hat{\boldsymbol{\theta}}_r(k) = [\hat{a}_1(k), \hat{\phi}_1(k), \dots, \hat{a}_n(k), \hat{\phi}_n(k)]^T. \quad (22)$$

The steps of computing the characteristic parameter estimate are as follows.

1. To initialize: let $k = 0$ and preset recursive length L ; let $\hat{\boldsymbol{\theta}}(0)$ be an arbitrary small vector; let $p_0 = 10^6$.
2. Collect the measured data $y(k)$.
3. Compute $\boldsymbol{\varphi}(k)$ using (20); compute $\mathbf{P}(k)$ using (19).
4. Compute $\mathbf{L}(k)$ using (18).
5. Update the parameter estimate $\hat{\boldsymbol{\theta}}(k)$ using (17) and obtain the characteristic parameter estimate $\hat{\boldsymbol{\theta}}_r(k)$ using (21)–(22), if $k = L$, terminate the recursive procedure and obtain the parameter estimate $\hat{\boldsymbol{\theta}}_r(k)$; otherwise, go to Step 2.

5 Illustrative Examples

Example 1 Consider the combination sine signals with four different frequencies,

$$y(t) = 1.8 \sin(0.07t + 0.95) + 2.9 \sin(0.5t + 0.8) + 4 \sin(2t + 0.76) + 2.5 \sin(1.6t + 1.1),$$

where $a_1 = 1.8$, $a_2 = 2.9$, $a_3 = 4$, $a_4 = 2.5$, $\phi_1 = 0.95$, $\phi_2 = 0.8$, $\phi_3 = 0.76$, $\phi_4 = 1.1$ are the true values of the parameters to be estimated and $\omega_1 = 0.07$ rad/s, $\omega_2 = 0.5$ rad/s, $\omega_3 = 2$ rad/s, $\omega_4 = 1.6$ rad/s are the known angular frequency.

Case 1 The MISG method simulation.

Use the proposed MISG algorithm to estimate the characteristic parameters of the combination sine signals in this example. In the simulation, the white noise sequence with zero mean and variance $\sigma^2 = 0.20^2$ is added to the signal. The sampling period is $h = 0.2$ s, and the data length is $L = 2000$. In order to test the performance of the MISG algorithm, three different innovation length data with $p = 1$, $p = 4$ and $p = 6$ are adopted. The parameter estimates and their estimation errors $\delta := \|\hat{\theta}(k) - \theta\|/\|\theta\|$ are listed in Table 1. The parameter estimation errors $\delta := \|\hat{\theta}(k) - \theta\|/\|\theta\|$ versus k are shown in Fig. 1. In addition, the estimated signal and the actual signal are compared for testing the estimation accuracy. The comparison results are shown in Fig. 2, where the dot-line denotes the estimated signal and the solid-line denotes the actual signal.

Case 2 The RLS method simulation.

Next, the proposed RLS algorithm is used to estimate the characteristic parameters. In the simulation, the white noise sequence with zero mean and variance $\sigma^2 = 0.10^2$, $\sigma^2 = 0.50^2$ are added, respectively, to the combination sine signal. In the simulation, the sampling period is $h = 1$ s and the data length is $L = 2000$. The parameter estimates and their estimation errors $\delta := \|\hat{\theta}(k) - \theta\|/\|\theta\|$ are shown in Table 2. The parameter estimation errors versus k are shown in Fig. 3.

For the purpose of testing the performance of the proposed RLS method, the estimated combination sine signal obtained by the RLS method and the actual combination sine signal are shown in Fig. 4.

Example 2 In this example, a period signal is provided to test the proposed parameter estimation method. Consider a period square wave with the following description,

$$f(t) = \begin{cases} A, & -\frac{T}{2} \leq t < 0, \\ -A, & 0 \leq t \leq \frac{T}{2}, \end{cases}$$

where $T = \frac{2\pi}{3}$ s, $A = 2$. The period square wave is shown in Fig. 5.

Using the Fourier expansion, this square wave can be expanded into the sum of odd harmonics. According to the Fourier expansion formula, the coefficients are, respectively

Table 1 The MSG parameter estimates and their estimation errors

p	k	a_1	a_2	a_3	a_4	ϕ_1	ϕ_2	ϕ_3	ϕ_4	δ (%)
1	100	2.32836	3.69068	4.59871	3.12327	1.06778	-1.02725	1.43986	2.52419	44.86674
	200	2.32210	3.67568	4.58352	3.11300	1.00956	-1.02409	1.40327	2.48824	44.16653
	500	2.29745	3.65673	4.56484	3.09331	0.98713	-1.04918	1.32503	2.45879	43.60084
	1000	2.28357	3.64158	4.55076	3.08024	0.96741	-1.06163	1.27192	2.43280	43.13828
	1500	2.27537	3.63295	4.54221	3.07176	0.95213	-1.06978	1.23684	2.41775	42.87553
	2000	2.26868	3.62685	4.53608	3.06615	0.94130	-1.07605	1.21595	2.40565	42.70004
4	100	2.23287	3.61642	4.71487	3.37101	1.36497	0.45419	2.15869	2.23099	38.42730
	200	2.20612	3.56192	4.65818	3.32740	1.14917	0.46370	2.03040	2.10349	34.74527
	500	2.10709	3.48519	4.58289	3.24951	1.06175	0.36138	1.71218	1.98609	29.47719
	1000	2.05222	3.42506	4.52728	3.19601	0.98250	0.31284	1.50140	1.88287	25.92727
	1500	2.01938	3.39088	4.49261	3.16215	0.91895	0.28343	1.36329	1.82227	23.88320
	2000	1.99266	3.36636	4.46835	3.14012	0.87660	0.25728	1.27908	1.77450	22.63846
6	100	1.86406	3.31522	4.52434	3.30687	1.45780	1.08330	2.05926	1.83465	31.39003
	200	1.82167	3.23888	4.44449	3.24050	1.15102	1.08898	1.86610	1.66358	26.04888
	500	1.67194	3.12205	4.32986	3.12419	1.02227	0.93346	1.38419	1.48515	17.38410
	1000	1.59029	3.03227	4.24715	3.04262	0.90272	0.86148	1.06597	1.33385	12.38772
	1500	1.54100	2.98137	4.19508	2.99201	0.80496	0.81998	0.86160	1.24191	10.43925
	2000	1.50086	2.94451	4.15871	2.95926	0.74187	0.78040	0.73449	1.17013	10.06083
True values										
		1.80000	2.90000	4.00000	2.50000	0.95000	0.80000	0.76000	1.10000	

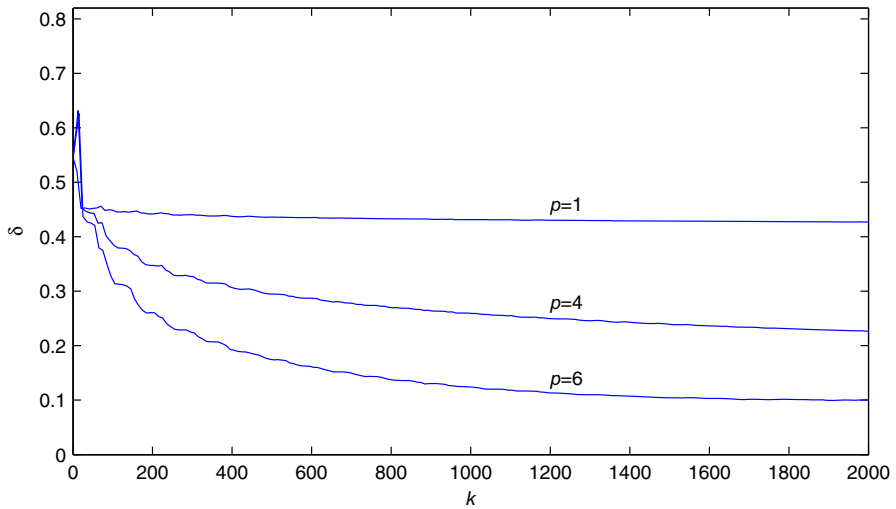


Fig. 1 The MISG estimation error δ versus k

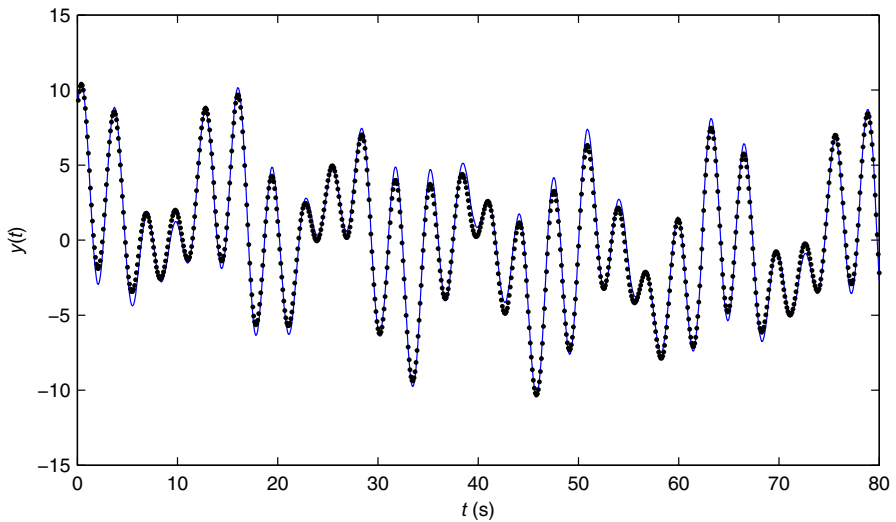


Fig. 2 The MISG fitting curves

$$\begin{aligned}
 a_i &= \frac{2}{T} \int_{-\frac{T}{2}}^0 -A \cos i\omega_0 t dt + \frac{2}{T} \int_0^{\frac{T}{2}} A \cos i\omega_0 t dt \\
 &= \frac{-2A}{Ti\omega_0} \sin i\omega_0 t \Big|_{-\frac{T}{2}}^0 + \frac{2A}{Ti\omega_0} \sin i\omega_0 t \Big|_0^{\frac{T}{2}} = 0, \\
 b_i &= \frac{2}{T} \int_{-\frac{T}{2}}^0 -A \sin i\omega_0 t dt + \frac{2}{T} \int_0^{\frac{T}{2}} A \sin i\omega_0 t dt
 \end{aligned}$$

Table 2 The RLS parameter estimates and their estimation errors

σ^2	k	a_1	a_2	a_3	a_4	ϕ_1	ϕ_2	ϕ_3	ϕ_4	δ (%)
0.20 ²	100	1.81292	2.88214	4.01877	2.51099	0.94731	0.79103	0.75548	1.12269	0.65179
	200	1.80360	2.87002	4.01002	2.48253	0.94945	0.78628	0.76026	1.11317	0.67165
	500	1.78622	2.88688	4.01268	2.48945	0.94680	0.79636	0.75765	1.10129	0.42264
	1000	1.78573	2.88650	3.99530	2.49662	0.95196	0.79805	0.76121	1.10207	0.34102
	1500	1.78643	2.88816	3.99611	2.49746	0.95218	0.80061	0.76171	1.10506	0.31929
	2000	1.79163	2.89189	3.99629	2.49857	0.95076	0.80182	0.76216	1.10433	0.21922
1.00 ²	100	1.86471	2.81311	4.09464	2.56755	0.93693	0.75404	0.73783	1.21118	3.29389
	200	1.81801	2.75576	4.05012	2.41712	0.94727	0.72851	0.76126	1.16767	3.28630
	500	1.73128	2.83477	4.06364	2.44730	0.93352	0.78145	0.74839	1.10656	2.11146
	1000	1.72870	2.83260	3.97657	2.48320	0.96011	0.79004	0.76607	1.11041	1.70349
	1500	1.73224	2.84082	3.98065	2.48796	0.96123	0.80312	0.76856	1.12539	1.59473
	2000	1.75814	2.85956	3.98165	2.49331	0.95385	0.80920	0.77083	1.12168	1.09406
True values		1.80000	2.90000	4.00000	2.50000	0.95000	0.80000	0.76000	1.10000	

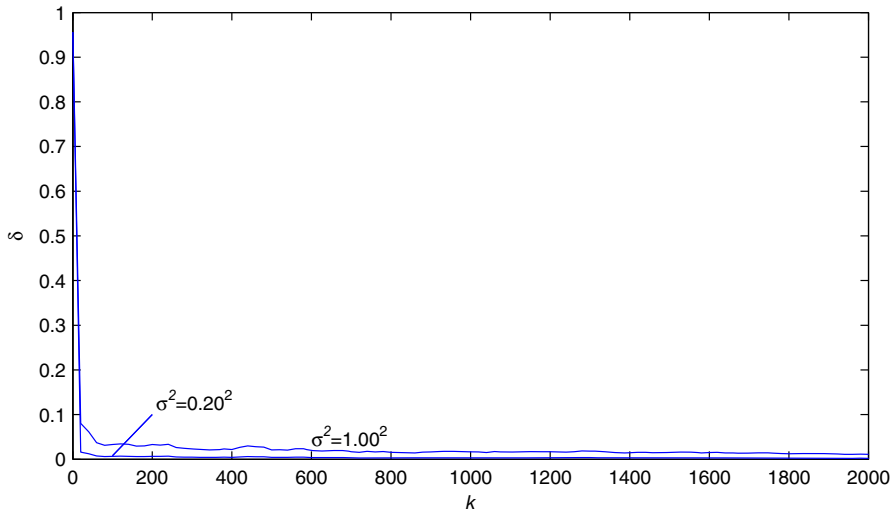


Fig. 3 The RLS estimation error δ versus k of Example 1

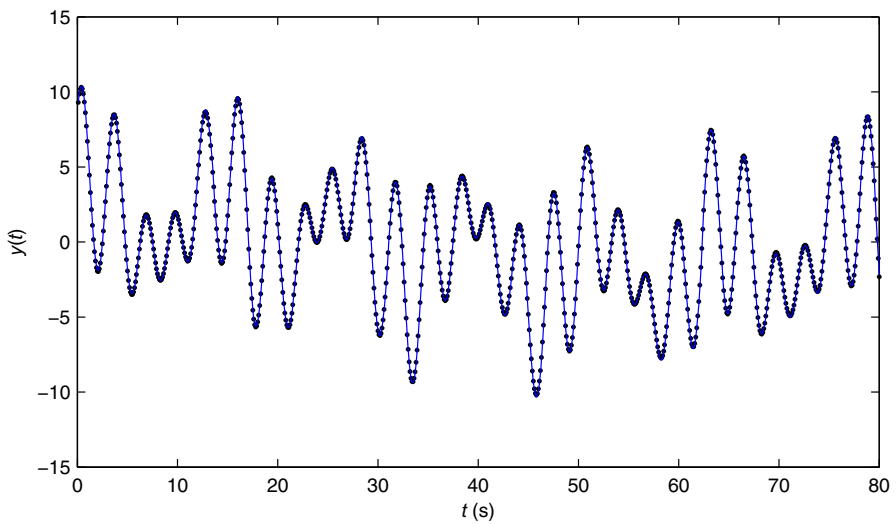


Fig. 4 The RLS fitting curves

$$= \frac{2A}{Ti\omega_0} \cos i\omega_0 t \Big|_{-\frac{T}{2}}^0 + \frac{-2A}{Ti\omega_0} \cos i\omega_0 t \Big|_0^{\frac{T}{2}} = \frac{A}{\pi i} (2 - 2 \cos i\pi).$$

As a result, when i is even number, $b_i = 0$; when i is odd number, $b_i = \frac{4A}{\pi i}$. Therefore, the Fourier expansion is given by

$$f(t) = \frac{4A}{\pi} \left(\sin \omega_0 t + \frac{1}{3} \sin 3\omega_0 t + \frac{1}{5} \sin 5\omega_0 t + \dots + \frac{1}{n} \sin n\omega_0 t + \dots \right).$$

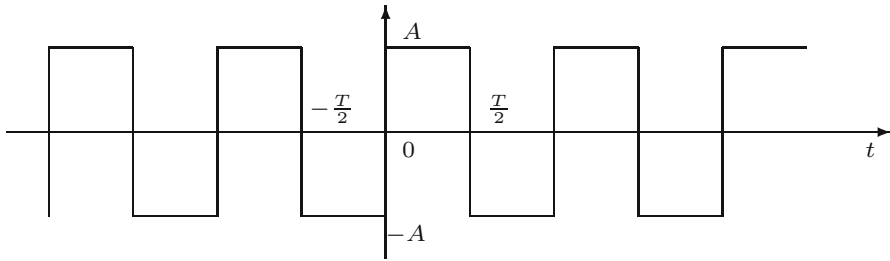


Fig. 5 The square signal wave

where ω_0 is the fundamental harmonic frequency and $\omega_0 = 2\pi/T$. The fundamental harmonic frequency equals the frequency of the square wave.

Taking the 1, 3, 5, 7 harmonics, we have

$$f(t) = \frac{4A}{\pi} \left(\sin \omega_0 t + \frac{1}{3} \sin 3\omega_0 t + \frac{1}{5} \sin 5\omega_0 t + \frac{1}{7} \sin 7\omega_0 t \right).$$

Let $a_1 := \frac{4A}{\pi}$, $a_2 := \frac{4A}{3\pi}$, $a_3 := \frac{4A}{5\pi}$ and $a_4 := \frac{4A}{7\pi}$, $f(t)$ becomes

$$f(t) = a_1 \sin \omega_0 t + a_2 \sin 3\omega_0 t + a_3 \sin 5\omega_0 t + a_4 \sin 7\omega_0 t.$$

In the above equation, the amplitudes are unknown, while the phases are zero. Using the proposed RLS parameter estimation method to estimate parameters a_1 , a_2 , a_3 and a_4 , the parameter estimates and their estimation errors $\delta := \|\hat{\theta}(k) - \theta\|/\|\theta\|$ are displayed in Table 3. The parameter estimation errors versus k are shown in Fig. 6.

Choosing the estimated parameters with $k = 2000$ and $\sigma^2 = 0.20^2$, we obtain the following function

$$f(t) = 2.54441 \sin 3t + 0.84968 \sin 9t + 0.51463 \sin 15t + 0.36396 \sin 21t.$$

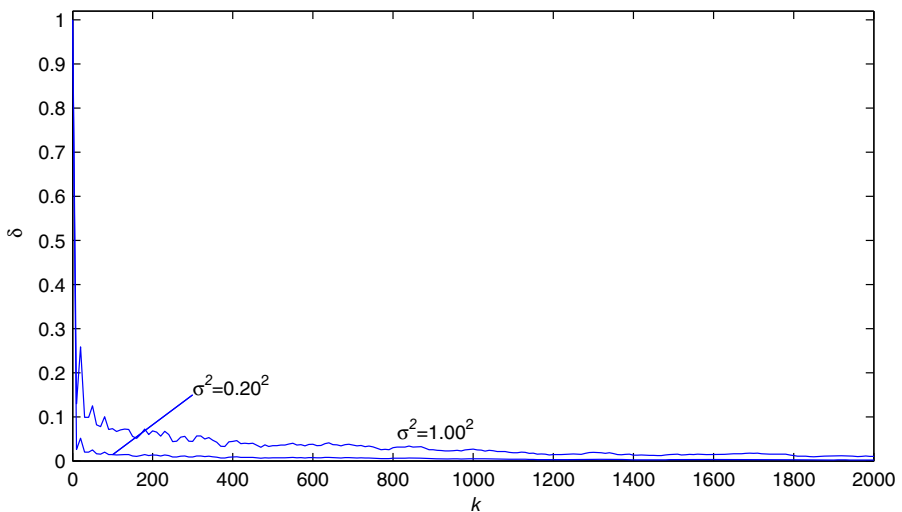
The estimated square wave and the original square wave are shown in Fig. 7.

From the simulation results, we can draw the following conclusions.

1. Table 1 and Fig. 1 show that the parameter estimates obtained by the MISG method become more accurate with the increasing in the innovation length p .
2. When $p = 1$, the MISG method degenerates into the SG method. From the last column in Table 1, it can be seen that the MISG method has higher accuracy than the SG method.
3. The parameter estimation errors given by the RLS algorithm become smaller with the increasing in k —see Figs. 3 and 6. The parameter estimation accuracy is related to the noise variance. The larger the noise variance is, the lower the parameter estimation accuracy is.
4. The fitting curve given by the RLS method is more closed to the actual signal curve than the fitting curve given by the MISG method—see Figs. 3 and 4. This means that the RLS algorithm has more effectiveness than the MISG method.

Table 3 The RLS parameter estimates and the estimation errors for the square wave

σ^2	k	a_1	a_2	a_3	a_4	δ (%)
0.20^2	100	2.53543	0.87808	0.52064	0.38648	1.46108
	200	2.52737	0.86779	0.51471	0.38923	1.35828
	500	2.52902	0.85563	0.51159	0.36131	0.69080
	1000	2.53433	0.84849	0.51734	0.36078	0.53985
	1500	2.54359	0.85424	0.50934	0.36927	0.29887
	2000	2.54441	0.84968	0.51463	0.36396	0.20979
1.00^2	100	2.49124	0.99509	0.56601	0.47729	7.30539
	200	2.45096	0.94367	0.53637	0.49101	6.79139
	500	2.45918	0.88286	0.52078	0.35144	3.45399
	1000	2.48574	0.84713	0.54950	0.34875	2.69923
	1500	2.53202	0.87591	0.50954	0.39124	1.49433
	2000	2.53616	0.85309	0.53595	0.36468	1.04895
True values		2.54648	0.84883	0.50930	0.36378	

**Fig. 6** The RLS estimation error δ versus k of Example 2

6 Conclusion

This paper considers the parameter estimation problems of the periodic signals based on the combination sine signals. According to the gradient searching, the SG parameter estimation algorithm for the combination sine signal is derived. On the basis of the SG algorithm, the MISG parameter estimation method is proposed for improving the estimation accuracy. Furthermore, the RLS parameter estimation algorithm is derived for enhancing the estimation accuracy by the function expansion. The simulation

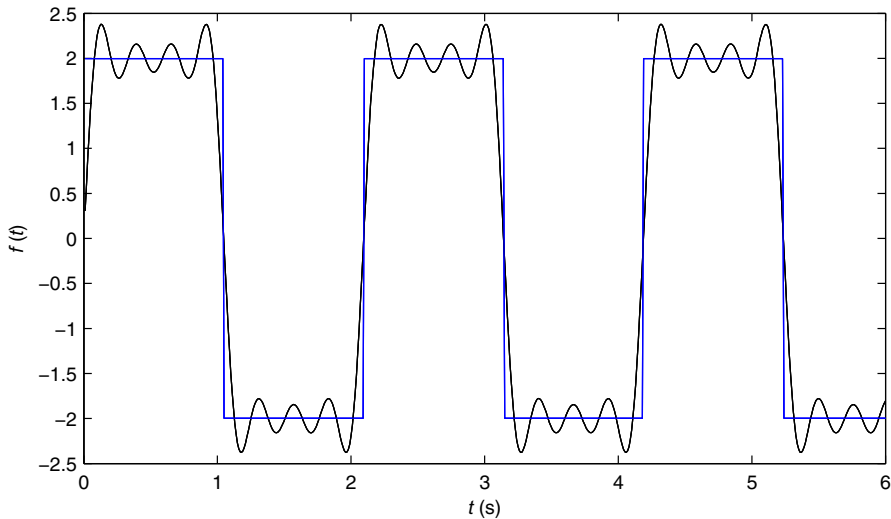


Fig. 7 The RLS square wave fitting curves of Example 2

results show that the MISG algorithm and the RLS method can estimate the signal parameters. Because the RLS algorithm is derived by means of the linear optimization while the MISG method is deduced by means of the nonlinear optimization principle, the RLS algorithm has higher accuracy and stabilization than the MISG method. The method used in this paper can be extended to analyze the convergence of the identification algorithms for linear or nonlinear control systems [6, 20, 21] and applied to hybrid switching-impulsive dynamical networks [18] and uncertain chaotic delayed nonlinear systems [17] or applied to other fields [4, 33, 34].

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 61273194) and Natural Science Fund for Colleges and Universities in Jiangsu Province (No. 12KJB120005).

References

1. N. Andrei, An adaptive conjugate gradient algorithm for large-scale unconstrained optimization. *J. Comput. Appl. Math.* **292**, 83–91 (2016)
2. D. Belega, D. Petri, Sine-wave parameter estimation by interpolated DFT method based on new cosine windows with high interference rejection capability. *Digit. Signal Process.* **33**, 60–70 (2014)
3. D. Belega, D. Petri, Accuracy analysis of the sine-wave parameters estimation by means of the windowed three-parameter sine-fit algorithm. *Digit. Signal Process.* **50**, 12–23 (2016)
4. X. Cao, D.Q. Zhu, S.X. Yang, Multi-AUV target search based on bioinspired neurodynamics model in 3-D underwater environments. *IEEE Trans. Neural Netw. Learn. Syst.* (2016). doi:[10.1109/TNNLS.2015.2482501](https://doi.org/10.1109/TNNLS.2015.2482501)
5. J. Chen, Y. Ren, G. Zeng, An improved multi-harmonic sine fitting algorithm based on Tabu search. *Measurement* **59**, 258–267 (2015)
6. Z.Z. Chu, D.Q. Zhu, S.X. Yang, Observer-based adaptive neural network trajectory tracking control for remotely operated Vehicle. *IEEE Trans. Neural Netw. Learn. Syst.* (2016). doi:[10.1109/TNNLS](https://doi.org/10.1109/TNNLS)
7. S. Deng, Z. Wan, A three-term conjugate gradient algorithm for large-scale unconstrained optimization problems. *Appl. Num. Math.* **92**, 70–81 (2015)

8. F. Ding, *System Identification-Performances Analysis for Identification Methods* (Science Press, Beijing, 2014)
9. F. Ding, X.M. Liu, Y. Gu, An auxiliary model based least squares algorithm for a dual-rate state space system with time-delay using the data filtering. *J. Franklin Inst.* **353**(2), 398–408 (2016)
10. F. Ding, P.X. Liu, G.J. Liu, Gradient based and least-squares based iterative identification methods for OE and OEMA systems. *Digit. Signal Process.* **20**(3), 664–677 (2010)
11. F. Ding, X.M. Liu, M.M. Liu, The recursive least squares identification algorithm for a class of Wiener nonlinear systems. *J. Franklin Inst.* **353**(7), 1518–1526 (2016)
12. F. Ding, X.M. Liu, X.Y. Ma, Kalman state filtering based least squares iterative parameter estimation for observer canonical state space systems using decomposition. *J. Comput. Appl. Math.* **301**, 135–143 (2016)
13. F. Ding, X.H. Wang, Q.J. Chen, Y.S. Xiao, Recursive least squares parameter estimation for a class of output nonlinear systems based on the model decomposition. *Circuits Syst. Signal Process.* **35**(9), 3323–3338 (2016)
14. J. Guo, Y.L. Zhao, C.Y. Sun, Y. Yu, Recursive identification of FIR systems with binary-valued outputs and communication channels. *Automatica* **60**, 165–172 (2015)
15. M. Jafari, M. Salimifard, M. Dehghani, Identification of multivariable nonlinear systems in the presence of colored noises using iterative hierarchical least squares algorithm. *ISA Trans.* **53**(4), 1243–1252 (2014)
16. A. Janot, P. Vandanjon, M. Gautier, A revised Durbin–Wu–Hausman test for industrial robot identification. *Control Eng. Pract.* **48**, 52–62 (2016)
17. Y. Ji, X.M. Liu, F. Ding, New criteria for the robust impulsive synchronization of uncertain chaotic delayed nonlinear systems. *Nonlinear Dyn.* **79**(1), 1–9 (2015)
18. Y. Ji, X.M. Liu, Unified synchronization criteria for hybrid switching-impulsive dynamical networks. *Circuits Syst. Signal Process.* **34**(5), 1499–1517 (2015)
19. X. Li, F. Ding, Signal modeling using the gradient search. *Appl. Math. Lett.* **26**(8), 807–813 (2013)
20. H. Li, Y. Shi, W. Yan, On neighbor information utilization in distributed receding horizon control for consensus-seeking. *IEEE Trans. Cybern.* (2016). doi:[10.1109/TCYB.2015.2459719](https://doi.org/10.1109/TCYB.2015.2459719)
21. H. Li, Y. Shi, W. Yan, Distributed receding horizon control of constrained nonlinear vehicle formations with guaranteed γ -gain stability. *Automatica* **68**, 148–154 (2016)
22. J. Li, Y.J. Zheng, Z.P. Lin, Recursive identification of time-varying systems: self-tuning and matrix RLS algorithms. *Syst. Control Lett.* **66**, 104–110 (2014)
23. Y.W. Mao, F. Ding, A novel data filtering based multi-innovation stochastic gradient algorithm for Hammerstein nonlinear systems. *Digit. Signal Process.* **46**, 215–225 (2015)
24. I. Necoara, V. Nedelcu, On linear convergence of a distributed dual gradient algorithm for linearly constrained separable convex problems. *Automatica* **55**, 209–216 (2015)
25. J. Pan, X.H. Yang, H.F. Cai, B.X. Mu, Image noise smoothing using a modified Kalman filter. *Neurocomputing* **173**, 1625–1629 (2016)
26. J. Vörös, Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities. *IEEE Trans. Autom. Control* **44**(11), 2145–2149 (1999)
27. D.Q. Wang, Hierarchical parameter estimation for a class of MIMO Hammerstein systems based on the reframed models. *Appl. Math. Lett.* **57**, 13–19 (2016)
28. D.Q. Wang, F. Ding, Parameter estimation algorithms for multivariable Hammerstein CARMA systems. *Inf. Sci.* **355–356**(10), 237–248 (2016)
29. Y.J. Wang, F. Ding, Novel data filtering based parameter identification for multiple-input multiple-output systems using the auxiliary model. *Automatica* **71**, 308–313 (2016)
30. Y.J. Wang, F. Ding, The filtering based iterative identification for multivariable systems. *IET Control Theory Appl.* **10**(8), 894–902 (2016)
31. Y.J. Wang, F. Ding, The auxiliary model based hierarchical gradient algorithms and convergence analysis using the filtering technique. *Signal Process.* **128**, 212–221 (2016)
32. Y.J. Wang, F. Ding, Recursive least squares algorithm and gradient algorithm for Hammerstein–Wiener systems using the data filtering. *Nonlinear Dyn.* **84**(2), 1045–1053 (2016)
33. T.Z. Wang, J. Qi, H. Xu et al., Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter. *ISA Trans.* **60**, 156–163 (2016)
34. T.Z. Wang, H. Wu, M.Q. Ni et al., An adaptive confidence limit for periodic non-steady conditions fault detection. *Mech. Syst. Signal Process.* **72–73**, 328–345 (2016)

35. D.Q. Wang, W. Zhang, Improved least squares identification algorithm for multivariable Hammerstein systems. *J. Franklin Inst.* **352**(11), 5292–5307 (2015)
36. L. Xu, A proportional differential control method for a time-delay system using the Taylor expansion approximation. *Appl. Math. Comput.* **236**, 391–399 (2014)
37. L. Xu, The damping iterative parameter identification method for dynamical systems based on the sine signal measurement. *Signal Process.* **120**, 660–667 (2016)
38. L. Xu, Application of the Newton iteration algorithm to the parameter estimation for dynamical systems. *J. Comput. Appl. Math.* **288**, 33–43 (2015)
39. L. Xu, L. Chen, W.L. Xiong, Parameter estimation and controller design for dynamic systems from the step responses based on the Newton iteration. *Nonlinear Dyn.* **79**(3), 2155–2163 (2015)
40. X.P. Xu, F. Wang, G.J. Liu, Identification of Hammerstein systems using key-term separation principle, auxiliary model and improved particle swarm optimisation algorithm. *IET Signal Process.* **7**(8), 766–773 (2013)
41. Y. Zhang, Unbiased identification of a class of multi-input single-output systems with correlated disturbances using bias compensation methods. *Math. Comput. Model.* **53**(9–10), 1810–1819 (2011)
42. G.Q. Zhang, X.K. Zhang, H.S. Pang, Multi-innovation auto-constructed least squares identification for 4 DOF ship manoeuvring modelling with full-scale trial data. *ISA Trans.* **58**, 186–195 (2015)
43. S.X. Zhao, F. Wang, H. Xu, J. Zhu, Multi-frequency identification method in signal processing. *Digit. Signal Process.* **19**(4), 555–566 (2009)